

What is claimed is:

1. A system for monitoring and controlling quality of service and availability of a discrete event system composed of a computer communications network having intermediate nodes, said system comprising :

an intermediate device, said intermediate device including memory for storing a traffic intensity surface therein, said intermediate device adapted to monitor traffic intensity by modeling the traffic intensity as a partial differential equation, said intermediate device sampling queue state information in intermediate nodes in said network and from this queue state information estimating the parameters of the partial differential equation model and solving the partial differential equation to produce the traffic intensity surface; said intermediate device configured to control traffic intensity by using several actuation mechanisms that include creating routes that minimize transit through congested regions of said network and downloading said routes to intermediate nodes for use in updating their forwarding tables, that include determining optimal times and locations for content caches to operate in said network and downloading said times and locations to cache content managers in said network, that include determining whether to correct bandwidth imbalances in said network either by buying and selling of short term bandwidth or by actuation of said network's topology and resources, including links and/or intermediate node capacities, and downloading said bandwidth actuations to bandwidth managers in the computer communications network; and

a set of traffic actuation devices, said devices including intermediate nodes responsible for relaying traffic between links in the computer communications network, including cache content managers responsible for deciding where to locate content caches in the computer communications network and when to have each content cache active, including bandwidth managers responsible for adding or deleting bandwidth in the computer communications network either temporarily or permanently.

2. A system as described in Claim 1 wherein said intermediate device is a computer system.

3. A system as described in Claim 1 wherein said intermediate device is an automatic traffic and quality of service control computer.

4. A system as described in Claim 1 wherein said intermediate device collects queue state information from intermediate nodes in the computer communications network that is being managed.

5. A system as described in Claim 1 wherein said intermediate device estimates the parameters of the partial differential equation that describes the traffic intensity dynamics of the discrete event system.

6. A system as described in Claim 1 wherein said intermediate device solves the partial differential equation that describes the traffic intensity dynamics of the computer communication network to yield the traffic intensity surface for the discrete event system.

7. A system as described in Claim 6 wherein said intermediate device generates optimal and suboptimal paths that are geodesics and non-geodesics, respectively, of the traffic intensity surface.

8. A system as described in Claim 7 wherein said intermediate device generates optimal and suboptimal routes by projecting the optimal and suboptimal paths on to the topology of the computer communications network.

9. A system as described in Claim 8 wherein said intermediate device actuates the traffic intensity by sending the optimal routes to intermediate nodes for the purpose of updating their forwarding tables.

10. A system as described in Claim 1 wherein said intermediate device determines the optimal times and locations for content caches to operate.

11. A system as described in Claim 10 wherein said intermediate device actuates the traffic intensity by downloading the optimal times and locations for content caches to operate to cache content managers in the computer communications network.

12. A system as described in Claim 1 wherein said intermediate device determines the times and locations of bandwidth imbalances in the computer communications network and their persistence.

13. A system as described in Claim 12 wherein said intermediate device actuates the traffic intensity by downloading the times and locations of the bandwidth imbalances to a bandwidth manager in the computer communications network.

14. An intermediate device for monitoring and controlling quality of service and availability of a discrete event system composed of a computer communications network having intermediate nodes, said intermediate device comprising :

memory for storing a traffic intensity surface therein, said intermediate device adapted to monitor traffic intensity by modeling the traffic intensity as a partial differential equation, said intermediate device sampling queue state information in intermediate nodes in said network and from this queue state information estimating the parameters of the partial differential equation model and solving the partial differential equation to produce the traffic intensity surface; said intermediate device adapted to control traffic intensity by using several actuation mechanisms that include creating routes that minimize transit through congested regions of said network and downloading said routes to intermediate nodes for use in updating their forwarding tables, that include determining optimal times and locations for content caches to operate in said network and downloading said times and locations to cache content managers in said network, that include determining whether to correct bandwidth imbalances in said network either by buying and selling of short term bandwidth or by actuation of said network's topology and resources, including links and/or intermediate node capacities, and downloading said bandwidth actuations to bandwidth managers in the computer communications network.

15. A system as described in Claim 14 wherein said intermediate device is a computer system.

16. A system as described in Claim 14 wherein said intermediate device is an automatic traffic and quality of service control computer.

17. A system as described in Claim 14 wherein said monitoring and controlling causes the quality of service and availability variables for the discrete event system to improve.

18. A method for monitoring and controlling the quality of service and availability variables for a discrete event system composed of a computer communications network and its traffic load, said method comprising the steps of:

(a) collecting queue state information from intermediate nodes in the computer communications network that is being managed and storing these in memory of an intermediate device;

(b) estimating the parameters of the partial differential equation that describes the traffic intensity dynamics of the discrete event system and storing these in memory of an intermediate device;

(c) solving the partial differential equation that describes the traffic intensity dynamics of the computer communication network to yield the traffic intensity surface for the discrete event system and storing this in memory of an intermediate device;

(d) generating optimal and suboptimal paths that are geodesics and non-geodesics, respectively, of the traffic intensity surface and storing these in memory of an intermediate device;

(e) generating optimal and suboptimal routes by projecting the optimal and suboptimal paths on to the topology of the computer communications network and storing these in memory of an intermediate device;

(f) actuating the traffic intensity by downloading the optimal routes to intermediate nodes for the purpose of updating their forwarding tables; and/or

(g) determining the optimal times and locations for content caches to operate and storing these in memory of an intermediate device;

(h) actuating the traffic intensity by downloading the optimal times and locations for content caches to operate to cache content managers in the computer communications network; and/or

(i) determining the times and locations of bandwidth imbalances in the computer communications network and their persistence and storing these in memory of an intermediate device; and

(j) actuating the traffic intensity by downloading the times and locations of the bandwidth imbalances to a bandwidth manager in the computer communications network.

19. A method as described in Claim 18 wherein step (a) comprises collecting said queue state information and storing them in memory of a computer system.

20. A method as described in Claim 18 wherein step (a) comprises collecting said queue state information and storing them in memory of an automatic traffic and quality of service control computer.

21. A method as described in Claim 18 wherein step (b) comprises estimating the parameters of the partial differential equation that describes the traffic intensity dynamics of the discrete event system and storing them in memory of a computer system.

22. A method as described in Claim 18 wherein step (b) comprises estimating the parameters of the partial differential equation that describes the traffic intensity dynamics of the discrete event system and storing them in memory of an automatic traffic and quality of service control computer.

23. A method as described in Claim 18 wherein step (c) comprises solving the partial differential equation that describes the traffic intensity dynamics of the computer communication network to yield the traffic intensity surface for the discrete event system and storing this in memory of a computer system.

24. A method as described in Claim 18 wherein step (c) comprises solving the partial differential equation that describes the traffic intensity dynamics of the computer communication network to yield the traffic intensity surface for the discrete event system and storing this in memory of an automatic traffic and quality of service control computer.

25. A method as described in Claim 18 wherein step (d) comprises generating optimal and suboptimal paths that are geodesics and non-geodesics, respectively, of the traffic intensity surface and storing these in memory of a computer system.

26. A method as described in Claim 18 wherein step (d) comprises generating optimal and suboptimal paths that are geodesics and non-geodesics, respectively, of the traffic intensity surface and storing these in memory of an automatic traffic and quality of service control computer.

27. A method as described in Claim 18 wherein step (e) comprises generating optimal and suboptimal routes by projecting the optimal and suboptimal paths on to the topology of the computer communications network and storing these in memory of a computer system.

28. A method as described in Claim 18 wherein step (e) comprises generating optimal and suboptimal routes by projecting the optimal and suboptimal paths on to the topology of the computer communications network and storing these in memory of an automatic traffic and quality of service control computer.

29. A method as described in Claim 18 wherein step (f) comprises actuating the traffic intensity by downloading the optimal routes to intermediate nodes for the purpose of updating their forwarding tables, wherein said updating of their forwarding tables causes the intermediate nodes to improve the computer communications network's quality of service and availability by routing traffic around congestion and regions of high utilization in the computer communication network.

30. A method as described in Claim 18 wherein step (f) comprises actuating the traffic intensity by downloading the optimal routes to intermediate nodes for the purpose of updating their forwarding tables, wherein said updating of their forwarding tables causes the intermediate nodes to improve the computer communications network's quality of service and availability by routing traffic over a large percentage of available links in the computer communication network.

31. A method as described in Claim 18 wherein step (g) comprises determining the optimal times and locations for content caches to operate and storing these in memory of a computer system.

32. A method as described in Claim 18 wherein step (g) comprises determining the optimal times and locations for content caches to operate and storing these in memory of an automatic traffic and quality of service control computer.

33. A method as described in Claim 18 wherein step (h) comprises actuating the traffic intensity by downloading the optimal times and locations for content caches to operate to cache content managers in the computer communications network, wherein said optimal cache times and

locations causes the cache content managers to improve the computer communications network's quality of service and availability by avoiding congested regions of the computer communications network, and to specify the optimal times when requests should be redirected to given cache locations.

34. A method as described in Claim 18 wherein step (i) comprises determining the times and locations of bandwidth imbalances in the computer communications network and their persistence and storing these in memory of a computer system.

35. A method as described in Claim 18 wherein step (i) comprises determining the times and locations of bandwidth imbalances in the computer communications network and their persistence and storing these in memory of an automatic traffic and quality of service control computer.

36. A method as described in Claim 18 wherein step (j) comprises actuating the traffic intensity by downloading the times and locations of the bandwidth imbalances to bandwidth managers in the computer communications network, wherein said times and locations cause the bandwidth managers to improve the computer communications network's quality of service and availability by correcting bandwidth imbalances in the computer communications network either by buying and selling of short term bandwidth or by actuation of said network's topology and resources, including links and/or intermediate node capacities.

37. In a computer system having a processor coupled to a bus, a computer system coupled to said bus and having stored therein a computer program that when executed by said processor causes said computer system to implement a method for managing quality of service and availability in a computer communications network, said method comprising the steps of:

(a) collecting queue state information from intermediate nodes in the computer communications network that is being managed and storing these in memory of an intermediate device;

(b) estimating the parameters of the partial differential equation that describes the traffic intensity dynamics of the discrete event system and storing these in memory of an intermediate device;

(c) solving the partial differential equation that describes the traffic intensity dynamics of the computer communication network to yield the traffic intensity surface for the discrete event system and storing this in memory of an intermediate device;

(d) generating optimal and suboptimal paths that are geodesics and non-geodesics, respectively, of the traffic intensity surface and storing these in memory of an intermediate device;

(e) generating optimal and suboptimal routes by projecting the optimal and suboptimal paths on to the topology of the computer communications network and storing these in memory of an intermediate device;

(f) actuating the traffic intensity by downloading the optimal routes to intermediate nodes for the purpose of updating their forwarding tables; and/or

(g) determining the optimal times and locations for content caches to operate and storing these in memory of an intermediate device;

(h) actuating the traffic intensity by downloading the optimal times and locations for content caches to operate to cache content managers in the computer communications network; and/or

(i) determining the times and locations of bandwidth imbalances in the computer communications network and their persistence and storing these in memory of an intermediate device; and

(j) actuating the traffic intensity by downloading the bandwidth imbalances to a bandwidth manager in the computer communications network.

38. A computer system as described above in Claim 37 wherein step (a) of said computer implemented method stored on said computer system comprises collecting said queue state information and storing it in memory of a computer system.

39. A system medium as described above in Claim 37 wherein step (a) of said computer implemented method stored on said computer system comprises collecting said queue state information and storing it in memory of an automatic traffic and quality of service control computer.

40. A computer system as described above in Claim 37 wherein step (b) of said computer implemented method stored on said computer system comprises estimating the parameters of the partial differential equation that describes the traffic intensity dynamics of the discrete event system and storing these in memory of a computer system.



41. A computer system as described above in Claim 37 wherein step (b) of said computer implemented method stored on said computer system comprises estimating the parameters of the partial differential equation that describes the traffic intensity dynamics of the discrete event system and storing these in memory of an automatic traffic and quality of service control computer.

42. A computer system as described above in Claim 37 wherein step (c) of said computer implemented method stored on said computer system comprises solving the partial differential equation that describes the traffic intensity dynamics of the computer communication network to yield the traffic intensity surface for the discrete event system and storing this in memory of a computer system.

43. A computer system as described above in Claim 37 wherein step (c) of said computer implemented method stored on said computer system comprises solving the partial differential equation that describes the traffic intensity dynamics of the computer communication network to yield the traffic intensity surface for the discrete event system and storing this in memory of an automatic traffic and quality of service control computer.

44. A computer system as described above in Claim 37 wherein step (d) of said computer implemented method stored on said computer system comprises generating optimal and suboptimal paths that are geodesics and non-geodesics, respectively, of the traffic intensity surface and storing these in memory of a computer system.

45. A computer system as described above in Claim 37 wherein step (d) of said computer implemented method stored on said computer system comprises generating optimal and suboptimal paths that are geodesics and non-geodesics, respectively, of the traffic intensity surface and storing these in memory of an automatic traffic and quality of service control computer.

46. A computer system as described above in Claim 37 wherein step (e) of said computer implemented method stored on said computer system comprises generating optimal and suboptimal routes by projecting the optimal and suboptimal paths on to the topology of the computer communications network and storing these in memory of a computer system.

47. A computer system as described above in Claim 37 wherein step (e) of said computer implemented method stored on said computer system comprises generating optimal and suboptimal routes by projecting the optimal and suboptimal paths on to the topology of the computer communications network and storing these in memory of an automatic traffic and quality of service control computer.

48. A computer system as described above in Claim 37 wherein step (f) of said computer implemented method stored on said computer system comprises actuating the traffic intensity by downloading the optimal routes to intermediate nodes for the purpose of updating their forwarding tables, wherein said updating of their forwarding tables causes the intermediate nodes to improve the computer communications network's quality of service and availability by routing traffic around congestion and regions of high utilization in the computer communication network.

49. A computer system as described above in Claim 37 wherein step (g) of said computer implemented method stored on said computer system comprises determining the optimal times and locations for content caches to operate and storing these in memory of a computer system.

50. A computer system as described above in Claim 37 wherein step (g) of said computer implemented method stored on said computer system comprises determining the optimal times and locations for content caches to operate and storing these in memory of an automatic traffic and quality of service control computer.

51. A computer system as described above in Claim 37 wherein step (h) of said computer implemented method stored on said computer system comprises actuating the traffic intensity by downloading the optimal times and locations for content caches to operate to cache content managers in the computer communications network, wherein said optimal cache times and locations cause the cache content managers to improve the computer communications network's quality of service and availability by avoiding congested regions of the computer communications network, and to specify the optimal times when requests should be redirected to given cache locations.

52. A computer system as described above in Claim 37 wherein step (i) of said computer implemented method stored on said computer system comprises determining the times and locations

of bandwidth imbalances in the computer communications network and their persistence and storing these in memory of a computer system.

53. A computer system as described above in Claim 37 wherein step (i) of said computer implemented method stored on said computer system comprises determining the times and locations of bandwidth imbalances in the computer communications network and their persistence and storing these in memory of an automatic traffic and quality of service control computer.

54. A computer system as described above in Claim 37 wherein step (j) of said computer implemented method stored on said computer system comprises actuating the traffic intensity by downloading the times and locations of the bandwidth imbalances to bandwidth managers in the computer communications network, wherein said times and locations cause the bandwidth managers to improve the computer communications network's quality of service and availability by correcting bandwidth imbalances in the computer communications network either by buying and selling of short term bandwidth or by actuation of said network's topology and resources, including links and/or intermediate node capacities.

55. A system for monitoring and controlling quality of service and availability in a communications comprising,  
a processor including memory for storing a traffic intensity surface,  
said processor configured to sample queue state information for selected intermediate nodes in said network to provide queue state information and processing said state information to produce said traffic intensity surface, and  
said processor being further configured to used said traffic intensity surface to control the traffic in the network.

56. A system as in Claim 55 in which said traffic control includes generating paths that route traffic around congestion and regions of high utilization.

57. A system as in Claim 55 in which the process is further configured to generate the optimal locations at which caches of replicated data may be located to avoid congested regions of the network and to specify the optimal times when requests should be redirected to given cache locations.

58. A system as in Claim 55 in which said traffic control includes generating paths that route traffic over a larger percentage of available links.

59. A system as in Claim 55 in which said processor is configured to provide traffic intensity and utilization state information to capacity planning tools that seek to adjust traffic and bandwidth.

60. A system as in Claim 55 in which said processor is additionally configured to determine when certain links and/or intermediate nodes of the network have surplus bandwidth or a bandwidth deficit and the level of persistence of such bandwidth imbalances and when such imbalances should be actuated using either a bandwidth trading tool which will remedy or purchase bandwidth to actuate bandwidth deficits and/or make available surplus capacity for resale to third party traffic, or when such imbalances should be actuated by actuating the network's topology and/or bandwidth.